# Solution Structure of the Carboxyl-Terminal Domain of RAP74 and NMR Characterization of the FCP1-Binding Sites of RAP74 and Human TFIIB[†,‡]

Bao D. Nguyen,[§] Hung-Ta Chen,[||] Michael S. Kobor,[⊥,#] Jack Greenblatt,[⊥,#] Pascale Legault,[§] and James G. Omichinski*[,§,||,@]

*Department of Biochemistry and Molecular Biology and Department of Chemistry, University of Georgia, Athens, Georgia 30602, Banting and Best Department of Medical Research and Department of Molecular and Medical Genetics, University of Toronto, Toronto, Ontario, Canada M5G 1L6, and Institute of Structural Biology, IBI-2, Forschungszentrum Jülich, 52424 Jülich, Germany*

*Received July 31, 2002; Revised Manuscript Received December 14, 2002*

ABSTRACT: FCP1 (TFIIF-associated CTD phosphatase) is the only known phosphatase specific for the phosphorylated CTD of RNAP II. The phosphatase activity of FCP1 is strongly enhanced by the carboxyl-terminal domain of RAP74 (cterRAP74, residues 436−517), and this stimulatory effect of TFIIF can be blocked by TFIIB. It has been shown that cterRAP74 and the core domain of hTFIIB (TFIIBc, residues 112−316) directly interact with the carboxyl-terminal domain of hFCP1 (cterFCP, residues 879−961), and these interactions may be responsible for the regulatory activities of TFIIF and TFIIB on FCP1. We have determined the NMR solution structure of human cterRAP74, and we have used NMR methods to map the cterFCP-binding sites for both cterRAP74 and human TFIIB. We show that cterFCP binds to a groove of cterRAP74 between α-helices H2 and H3, without affecting the secondary structure of cterRAP74. We also show that cterFCP binds to a groove of TFIIBc between α-helices D1 and E1 in the first cyclin repeat. We find that the cterFCP-binding site of TFIIBc is very similar to the binding site for the HSV transcriptional activator protein VP16 on the first cyclin repeat of TFIIBc. The cterFCP-binding sites of both RAP74 and TFIIBc form shallow grooves on the protein surface, and they are both rich in hydrophobic and positively charged amino acid residues. These results provide new information about the recognition of acidic-rich activation domains involved in transcriptional regulation, and provide insights into how TFIIF and TFIIB regulate the FCP1 phosphatase activity in vivo.

In human cells, most promoters of protein-coding genes transcribed by RNA polymerase II (RNAP II) contain the essential TATA box DNA sequence located approximately 30 base pairs upstream of the transcription start site. The assembly of RNAP II and the general transcription factors (GTF) required for basal transcription onto the DNA promoter are termed the transcription preinitiation complex (PIC) (for reviews, see refs *1* and *2*). The human PIC is comprised of more than 30 proteins, including RNAP II and general transcription factors TFIID, TFIIB, TFIIF, TFIIE, TFIIH, and, in many cases, TFIIA (*3*, *4*). Previous experiments have indicated that the PIC may form in vivo by promoter binding of a preassembled holoenzyme complex containing RNAP II and many, if not all, of the essential transcription factors needed for initiation (*5*). However, other experiments have demonstrated that the general transcription factors and RNAP II can be assembled in vitro in an extremely ordered fashion that is nucleated by binding of the TATA-binding protein (TBP) subunit of TFIID to the TATA box (*6*).

RNAP II is a multisubunit enzyme complex, which consists of two large subunits and numerous other smaller subunits (*1*). RNAP II differs from RNAP I and RNAP III in that its largest subunit possesses a distinct repeating heptapeptide unit with the Tyr-Ser-Pro-Thr-Ser-Pro-Ser consensus sequence that defines the carboxyl-terminal domain (CTD). This repeating heptapeptide unit varies in length depending on the organism. In mammalian cells, the CTD contains 52 units, while the yeast CTD has 26−27 units (*7*, *8*). In all cases, the CTD repeat is essential for survival of the organism. Interestingly, in addition to its roles in transcriptional regulation described below, the CTD has been implicated in pre-mRNA processing, including splicing, 5′-capping, and 3′-end polyadenylation (*5*, *9*, *10*).

RNAP II enters the initiation complex with its CTD in the unphosphorylated form (RNAP IIA), and the CTD becomes extensively phosphorylated (RNAP IIO) within the heptapeptide repeat during the early stages of the transcription process (*11*−*13*). The general transcription factor TFIIH

[†] This work was supported by National Institutes of Health Grant RO1 GM60298-01 (J.G.O. and P.L.) and by a grant from the Canadian Institutes of Health Research (J.G.).

[‡] Coordinates deposited as PDB entry HPUB.

* To whom correspondence should be addressed: Department of Biochemistry and Molecular Biology, Life Sciences Building, University of Georgia, Athens, GA 30602. Phone: (706) 542-6116. E-mail: jim@bmb.uga.edu.

[§] Department of Biochemistry and Molecular Biology, University of Georgia.

[||] Department of Chemistry, University of Georgia.

[⊥] Banting and Best Department of Medical Research, University of Toronto.

[#] Department of Molecular and Medical Genetics, University of Toronto.

[@] Forschungszentrum Jülich.

10.1021/bi0265473 CCC: $25.00 
Published on Web 01/24/2003

plays a key role in the transition from the initiation phase to the elongation phase of transcription by virtue of its ability to phosphorylate the CTD of RNAP II (*14*−*16*). In addition to TFIIH, many kinases have been implicated in the phosphorylation of the CTD (*14*−*18*). This phosphorylation of the CTD occurs primarily at serine 2 and serine 5 of the heptapeptide repeats, and this enables RNAP II to progress from the initiation complex to a stable elongation complex (*19*). The CTD remains extensively phosphorylated throughout the elongation phase. For the same RNAP II to be recycled back to the initiation complex, the CTD must be in the unphosphorylated form (RNAP IIA); therefore, dephosphorylation of the CTD by a phosphatase(s) is essential for generation of a form of the RNAP II that is capable of reinitiating transcription (*20*).

A unique phosphatase specific for the phosphorylated CTD of RNAP II was originally purified from HeLa cell extracts (*21*) and subsequently from *Saccharomyces cerevisiae* (*22*). From experiments with the CTD phosphatase purified from HeLa cells, it was found that the evolutionarily conserved carboxyl-terminal domain of the large subunit of TFIIF (RAP74 in humans) strongly stimulates the CTD phosphatase activity (*23*). Conversely, TFIIB was also shown to play a critical role in regulating the CTD phosphatase activity through its ability to block the stimulatory effect produced by RAP74 (*23*). Using a yeast two-hybrid screen to identify proteins that bind to the conserved carboxyl-terminal domain of RAP74 (cterRAP74, residues 436−517), partial cDNAs encoding a human CTD phosphatase (hFCP1) were isolated (*24*). Subsequent studies have identified the complete cDNA of hFCP1 (*25*) and the complete cDNA of the homologous yeast FCP1 (yFCP1) (*24*, *26*).

Three domains have been defined on the basis of the sequence similarity between hFCP1 and yFCP1 (*24*−*26*). The amino-terminal domain contains the catalytic portion of the protein and was originally termed the FCP1 homology domain. FCP1 is a member of a new class of small molecule phosphotransferases and phosphohydrolases that contain a conserved $\psi\psi\psi$DXDXT/V$\psi$ ($\psi$ is hydrophobic residue) catalytic motif (*25*, *27*−*29*). For previously characterized members of this family of phosphotransferases, the first aspartic acid within the DXDXT/V motif is phosphorylated in vivo, and this phosphorylated form of the enzyme may be the main intermediate in the phosphoryl transfer reaction (*27*). Mutations of the first aspartic acid within the DXDXT motif abolish the catalytic activity for several members of this family in vitro (*27*, *28*), and yeast strains carrying these mutations in yFCP1 are not viable (*29*). The remaining two domains of FCP1 appear to be involved in the regulation of the catalytic domain. The central domain contains a BRCT homology domain related to the breast cancer tumor suppressor protein BRCA1 and found in many proteins involved in DNA repair, recombination, and cell cycle control (*30*). The carboxyl-terminal domain is highly acidic, and its sequence resembles that of the acidic transactivation domain present in many regulatory transcription factors (*24*, *26*). Recently, it has been shown that this domain can function as a transactivation domain in vitro (*31*) or in vivo (J. Archambault and J. Greenblatt, unpublished data).

Binding experiments using yeast and human protein fragments have shown that the highly acidic carboxyl-terminal domain of FCP1 (cterFCP, residues 879−961 in humans) directly interacts with the conserved carboxyl-terminal domain of the large subunit of TFIIF (cterRAP74, residues 436−517 in humans) (*24*, *26*). In addition, it has also been shown that yeast cterFCP (residues 667−732) interacts with the first cyclin repeat (residues 101−240 in yeast) of the core domain of yeast TFIIB (TFIIBc) (*32*). The binding to FCP1 involves direct interactions with residues in a homologous stretch of eight amino acids found in TFIIF and TFIIB (*32*). In both TFIIF and TFIIB, these amino acid residues are part of an α-helix, and point mutations of some of these amino acid residues significantly alter FCP1 function (*32*−*34*).

In this paper, we describe the NMR solution structure of free human cterRAP74 and compare it with the crystal structure of zinc-bound human cterRAP74 (*34*). Then, we use $^{1}H-^{15}N$ HSQC experiments of $^{15}N$-labeled cterRAP74, free and bound to the unlabeled human cterFCP, to map the cterFCP-binding site on the solution structure of cterRAP74. In addition, using a complex of $^{13}C$- and $^{15}N$-labeled cterRAP74 bound to unlabeled cterFCP, we assign the backbone resonances of cterRAP74 in the complex with cterFCP. From a chemical shift index (CSI) analysis of the $H_\alpha$, $C_\alpha$, $C_\beta$, and C′ chemical shifts (*35*−*37*), we find that the secondary structure of cterRAP74 in the complex is very similar to that of free cterRAP74. We also use $^{1}H-^{15}N$ HSQC experiments of $^{15}N$-labeled human TFIIBc (residues 112−316), free and bound to the unlabeled human cterFCP, to map the cterFCP-binding site on the solution structure of TFIIBc (*33*). We find that the cterFCP-binding site of TFIIBc is very similar to the binding site for the herpes simplex virus (HSV) transcriptional activator protein VP16 on the first cyclin repeat of TFIIBc. Finally, we compare the cterFCP-binding sites of cterRAP74 and TFIIBc and find that they both form shallow grooves on the protein surface, which are created by similar secondary structures and display a comparable distribution of hydrophobic and basic amino acid residues. Our results provide new structural information about the recognition of acid-rich activation domains involved in transcriptional regulation, and provide insights into how TFIIF and TFIIB regulate the activity of FCP1, the only known CTD phosphatase.

## MATERIALS AND METHODS

*Protein Expression and Purification. (1) cterRAP74.* The cterFCP-binding domain of RAP74 (cterRAP74, amino acids 436−517) was expressed from a GST-2T vector (Pharmacia) as a GST fusion protein in *Escherichia coli* host strain BL21-(DE3) (*24*). Uniform (>98%) $^{15}N$ and $^{15}N/^{13}C$ labeling was carried out by growing the cells in a modified minimal medium containing $^{15}NH_4Cl$ and [$^{13}C_6$]glucose as the sole nitrogen and carbon sources, respectively. The cells were grown at 37 °C, and protein expression was induced for 4 h with 0.5 mM isopropyl D-thiogalactoside (IPTG). The cells (∼20 g) were harvested, suspended in binding buffer [25 mM Tris buffer (pH 7.4), 1 M NaCl, 1 mM DTT, and 1 mM EDTA], lysed by passage through a French press, and centrifuged at 105000*g* for 1 h. The supernatant was incubated for 2 h with 15 mL of glutathione−Sepharose resin (Pharmacia) at 4 °C. Following incubation, the resin was washed with binding buffer followed by equilibration with phosphate-buffered saline (PBS) containing 1 mM DTT. The

resin containing the bound fusion protein was then incubated overnight at room temperature with 100 units of thrombin (Calbiochem) to cut the cterRAP74 from the GST. Following thrombin cleavage, the supernatant was dialyzed into buffer A [20 mM phosphate buffer (pH 6.5), 1 mM DTT, and 1 mM EDTA] and then applied to an S-Sepharose Fast Flow (Pharmacia) column (100 mL bed volume) equilibrated with buffer A. The cterRAP74 was eluted from the column using a gradient (from 0 to 100% over 700 mL) of buffer B (20 mM phosphate, 1 mM DTT, 1 mM EDTA, and 1 M NaCl). The pooled fractions containing the purified cterRAP74 were then dialyzed into buffer A and concentrated for NMR analysis.

*(2) cterFCP*. The RAP74-binding domain of FCP (cterFCP, amino acids 879−961) was expressed from a GST-3X vector (Pharmacia) as a GST fusion protein in *E. coli* host strain BL21(DE3) (*24*). The cells were grown at 37 °C, and protein expression was induced for 4 h with 0.5 mM isopropyl D-thiogalactoside (IPTG). The cells (∼20 g) were harvested, resuspended in binding buffer, lysed by passage through a French press, and centrifuged at 105000*g* for 1 h. The supernatant was incubated for 2 h with 15 mL of glutathione−Sepharose resin (Pharmacia) at 4 °C. Following incubation, the resin was washed with binding buffer followed by equilibration with factor Xa cutting buffer [50 mM Tris (pH 8.0), 100 mM NaCl, and 1 mM $CaCl_2$]. The resin containing the bound fusion protein was then incubated overnight at room temperature with 100 units of factor Xa (Calbiochem) to cut the cterFCP from the GST. Following factor Xa cleavage, the supernatant was dialyzed into buffer A and then applied to a Q-Sepharose Fast Flow (Pharmacia) column (100 mL bed volume) equilibrated with buffer A. The cterFCP was eluted from the column using a gradient (from 0 to 100% over 700 mL) of buffer B. The pooled fractions containing the purified cterFCP were then dialyzed into buffer A and concentrated for NMR analysis.

*(3) TFIIBc*. The core domain of human TFIIB (TFIIBc, amino acids 1−3 and 112−316) was expressed from a pet3a vector in *E. coli* host strain BL21(DE3) (the clone was kindly provided by D. Reinberg) (*38*). Uniform (>98%) $^{15}N$ labeling was carried out by growing the cells in a modified minimal medium containing $^{15}NH_4Cl$ as the sole nitrogen source. The protein was expressed and purified using a published protocol (*38*). The purified protein was dialyzed into buffer A and concentrated for NMR analysis.

*Sample Preparation for NMR Spectroscopy*. The NMR samples were as follows. (1) For studies of the free cterRAP74, the sample consisted of 1 mM cterRAP74 (unlabeled, $^{15}N$-labeled, and $^{15}N$- and $^{13}C$-labeled) in 20 mM sodium phosphate (pH 6.5) and 1 mM EDTA (100% $D_2O$ or a 90% $H_2O$/10% $D_2O$ mixture). (2) For the cterRAP74−cterFCP complex, the sample consisted of 1 mM cterRAP74 ($^{15}N$-labeled and $^{15}N$- and $^{13}C$-labeled) and 1 mM unlabeled cterFCP in 20 mM sodium phosphate (pH 6.5) and 1 mM EDTA (90% $H_2O$/10% $D_2O$ mixture). (3) For studies of the free TFIIBc, the sample consisted of 0.75 mM $^{15}N$-labeled TFIIBc in 20 mM sodium phosphate (pH 6.5), 5 mM DTT, and 1 mM EDTA (90% $H_2O$/10% $D_2O$ mixture). (4) For the TFIIBc−cterFCP complex, the sample consisted of 0.75 mM $^{15}N$-labeled TFIIBc and 0.75 mM unlabeled cterFCP in 20 mM sodium phosphate (pH 6.5), 5 mM DTT, and 1 mM EDTA (90% $H_2O$/10% $D_2O$ mixture).

*NMR Spectroscopy*. NMR spectra were collected at 27 °C using Varian Inova Unity 500, 600, and 800 MHz NMR spectrometers equipped with a z pulsed-field gradient unit and triple-resonance probes. The backbone and aliphatic side chain signals of cterRAP74 ($^1H$, $^{15}N$, and $^{13}C$) were assigned using a combination of experiments, including two-dimensional (2D) $^1H-^{15}N$ HSQC (*39*), 2D $^1H-^{13}C$ CT-HSQC, three-dimensional (3D) $^{15}N$-edited TOCSY-HSQC (*40*), 3D HNCACB (*41*−*43*), 3D (HB)CBCA(CO)NNH (*42*, *43*), 3D HNCO (*44*), 3D H(CCO)NNH-TOCSY, 3D C(CO)NNH-TOSCY, and 4D HC(CO)NNH-TOCSY experiments (*43*, *45*). The aromatic $^1H$ and $^{13}C$ signals were assigned using a combination of 2D $^1H-^1H$ NOESY, 2D DQF-COSY, and 3D $^{13}C$-edited HMQC-NOESY experiments (*46*). Distance restraints were obtained from 2D $^1H-^1H$ NOESY (with mixing times of 50 and 150 ms at 500 MHz) (*47*), 3D $^{15}N$-edited NOESY-HSQC (with a mixing time of 150 ms at 600 MHz) (*40*), 3D $^{15}N$- and $^{13}C$-edited NOESY-HSQC (with mixing times of 50 and 150 ms at 600 MHz) (*48*), and 3D $^{13}C$-edited HMQC-NOESY (with a mixing time of 75 ms at 800 MHz) experiments (*46*). Chemical shifts of all proton, carbon, and nitrogen resonances were referenced externally to that of DSS at 0 ppm (*49*, *50*). The NMR data were processed with NMRPipe/NMRDraw (*51*) and analyzed with PIPP (*52*).

*Structure Calculation.* The NOE restraints were subdivided into four distance ranges: strong (1.8−2.7 Å), medium (1.8−3.3 Å), weak (1.8−5.0 Å), and very weak (1.8−6.0 Å) by comparison with NOEs involving protons separated by known distances within the α-helical elements (residues 456−465, 470−476, and 487−500). An extra 0.2 Å was added to the upper distance limit for NOE restraints in the medium and strong distance ranges that involved NH protons, while 0.5 Å was added to the upper distance limit for restraints involving methyl protons. A total of 959 NOE-derived distance restraints were obtained for structure calculation, including 586 short-range (361 intraresidue and 225 sequential), 189 medium-range, and 184 long-range distance restraints. Residues 436−451 at the amino terminus of the cterRAP74 did not display any medium-range or long-range NOEs, and these residues were therefore not included in the structure calculation. Backbone dihedral angle restraints ($\phi$ and $\psi$ angles) were obtained from analysis of $^1H_\alpha$, $^1H_N$, $^{13}C_\alpha$, $^{13}C_\beta$, $^{13}C'$, and $^{15}N$ chemical shifts using the program TALOS (*53*). The dihedral angle derived from TALOS is expressed as $\phi \pm$ error (or $\psi \pm$ error), where the value of the error used for structure calculations (from 15° to 40°) is twice that obtained from the TALOS analysis. A total of 104 TALOS-derived dihedral angle ($\phi$ and $\psi$) restraints were used for structure calculation.

Structures were calculated using the torsion angle molecular dynamics (TAD) protocol of CNS (*54*) starting from one extended structure with standard geometry. The protocol employed the conformational database potential derived from structure databases (*55*, *56*). From a total of 67 calculated structures, 50 of them satisfied all the experimental constraints with no NOE violation greater than 0.2 Å and no TALOS-derived dihedral angle ($\phi$ and $\psi$) restraints greater than 5°. From this set of 50 structures, the 20 structures with the lowest energy were selected for further analysis. In addition to standard protein covalent geometry, the force field included a van der Waals repulsive term, but neither

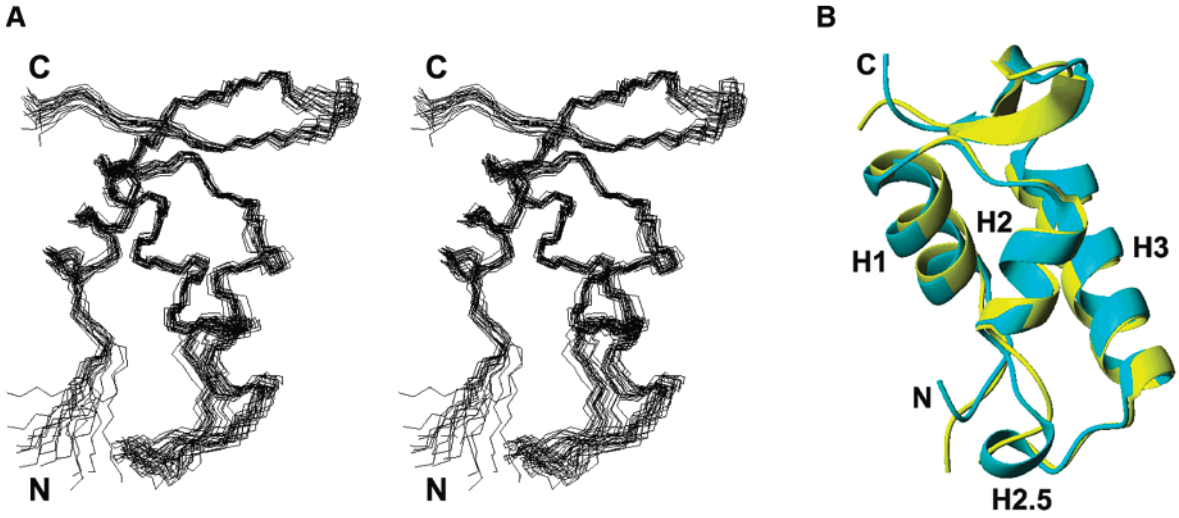


FIGURE 1: (A) Stereoview of the 20 lowest-energy NMR structures showing the backbone trace (N, $C_\alpha$, and C′ atoms) of residues 451−517 of the human cterRAP74. (B) Superposition of the NMR structure (yellow) of the free cterRAP74 and the X-ray crystal structure (cyan) of the zinc-bound cterRAP74 showing the backbone atoms (N, $C_\alpha$, and C′). The superposition is for heavy atoms of residues 451−517.

hydrogen bonding nor electrostatic terms were used at any stage of the structure calculations. The quality of the structures was analyzed using PROCHECK-NMR (*57*) and MOLMOL (*58*). All figures representing structures were generated using MOLMOL (*58*) and Molscript (*59*).

## RESULTS

*Resonance Assignments and Solution Structure of the Free cterRAP74*. All $^1$H, $^{15}$N, and $^{13}$C backbone and side chain resonances of cterRAP74 were assigned using a combination of 2D, 3D, and four-dimensional (4D) NMR experiments as described in detail in Materials and Methods. A total of 959 NOE-derived distance restraints were assigned from 2D $^1$H−$^1$H NOESY (*47*), 3D $^{13}$C-edited NOESY (*46*), 3D $^{15}$N-edited NOESY (*40*), and 3D $^{13}$C- and $^{15}$N-edited NOESY spectra (*48*). These restraints were supplemented with 104 TALOS-derived backbone dihedral angle ($\phi$ and $\psi$) restraints based on the analysis of the $^1H_\alpha$, $^1H_N$, $^{13}C_\alpha$, $^{13}C_\beta$, $^{13}C'$, and $^{15}$N chemical shifts (*53*). The three-dimensional structures of the cterRAP74 were calculated using the torsion angle molecular dynamics protocol of CNS (*54*) starting from an extended polypeptide structure of cterRAP74 residues 451−517. Due to the absence of medium-range and long-range NOEs for amino acids 436−450, these amino-terminal residues of the cterRAP74 were not included in the structure calculation. From the 67 calculated structures, a total of 50 were accepted because they have no distance restraint violation greater than 0.2 Å and no dihedral angle restraint ($\phi$ and $\psi$) violation greater than 5°. From these accepted structures, the 20 with the lowest energy were selected for further analysis (Table 1). From these 20 structures, an average structure was calculated and then minimized against the experimental restraints.

A backbone superposition of these 20 structures (Figure 1A) indicates that the structure of the cterRAP74 is precisely defined by the NMR data, except for one disordered loop (residues 477−486), which is devoid of medium- or long-range NOEs and flexible on the basis of $^{15}$N−$^1$H heteronuclear NOE data (Supporting Information). For the well-defined regions of the cterRAP74 (residues 456−476 and 487−514), the root-mean-square deviation (rmsd) from the minimized average structure for the ensemble of 20 structures is 0.53 ± 0.08 Å for the backbone atoms and 1.16 ± 0.10 Å for all heavy atoms. On the basis of the Ramachandran plot analysis in PROCHECK-NMR, 91.8% of $\phi$ and $\psi$ dihedral angles lie within the most favored region of the plot, while 7.6% lie in the additional allowed region. In this NMR structure, the cterRAP74 folds into three α-helices and one antiparallel $\beta$-sheet, with an $\alpha\alpha\alpha\beta\beta$ fold. The α-helices are from residue 456 to 465 (H1), from residue 470 to 476 (H2), and from residue 487 to 500 (H3). There is a four-residue loop (L1) between H1 and H2 and a 10-residue loop (L2)

Table 1: Structural Statistics for 20 Structures

| | |
|---|---|
| restraints used in the structure calculation[a] | |
| no. of distance restraints | |
| intraresidue | 361 |
| interresidue sequential ($i$, $i+1$) | 225 |
| interresidue medium-range ($1 < \|i-j\| < 5$) | 189 |
| long-range ($\|i-j\| \geq 5$) | 184 |
| total | 959 |
| rms deviations from distance restraints (Å) | 0.0033 ± 0.0014 |
| no. of torsion angle restraints ($\phi$ and $\psi$)[b] | 104 |
| rms deviations from dihedral restraints (deg) | 0.027 ± 0.009 |
| geometric statistics | |
| rms deviations from idealized geometry | |
| bonds (Å) | 0.00070 ± 0.00004 |
| angles (deg) | 0.3103 ± 0.0014 |
| impropers (deg) | 0.0970 ± 0.0017 |
| coordinate precision | |
| rms deviations from the minimized average structure[c] (Å) | |
| backbone atoms | 0.53 ± 0.08 |
| all heavy atoms | 1.16 ± 0.10 |
| Ramachandran analysis[d] | |
| most favored region (%) | 91.8 ± 2.1 |
| additionally allowed region (%) | 7.6 ± 2.0 |
| generously allowed region (%) | 0.6 ± 1.1 |
| disallowed regions (%) | 0 |

[a] None of these structures exhibited distance violations of >0.2 Å or dihedral angles of >5°. [b] Backbone dihedral angle restraints were generated with TALOS (*53*) based on analysis of $^1H_\alpha$, $^1H_N$, $^{13}C_\alpha$, $^{13}C_\beta$, $^{13}$C, and $^{15}$N chemical shifts. [c] The coordinate precision is defined as average atomic rms deviations between the 20 structures and the minimized average structure. The reported values are for residues 456−476 and 486−516 for the cterRAP74. [d] PROCHECK-NMR (*57*) was used to assess the quality of the structures.

between H2 and H3. The $\beta$-sheet is located at the carboxyl terminus of the domain, and is formed by residues 503−507 for the first strand (S1) and residues 510−514 (S2) for the second strand. The α-helices and the antiparallel $\beta$-sheet interact via the packing of the second strand of the $\beta$-sheet against loop L1 between helices H1 and H2. More precisely, residues F513 and L515 from the second strand of the $\beta$-sheet pack against hydrophobic residues of the three α-helices to form a hydrophobic core, which may be responsible for the stability of this domain in solution.

*Comparison of the Solution Structure of the Free cterRAP74 with the X-ray Structure of the Zinc-Bound cterRAP74.* A crystal structure of the zinc-bound cterRAP74 domain (residues 451−517) refined at 1.02 Å resolution has been reported recently (*34*). A backbone superposition between this structure and the minimized average NMR structure shows that the overall fold of the zinc-bound cterRAP74 in the crystal is very similar to that of the free cterRAP74 in solution (Figure 1B). An rmsd of 0.82 Å was calculated for the backbone atoms (N, Cα, and C′ for residues 456−476 and 487−514) of these two structures. There is only a subtle difference between these two structures, which is that residues 479−482 of L2 form a short α-helix in the crystal structure, while these residues are disordered in the NMR solution structure. Formation of a stable short helix in solution would give rise to characteristic backbone chemical shifts and NOEs ($H_{Ni}-H_{Ni+1}$ and $H_{\alpha i}-H_{\beta i+3}$), and these are not observed in the NMR spectra of the cterRAP74. In addition, $^{15}N-^{1}H$ heteronuclear NOE data indicate that this region is flexible in the NMR structure (Supporting Information).

To test if zinc is responsible for the difference in the two structures, we have obtained a $^{1}H-^{15}N$ HSQC spectrum of the cterRAP74 in the presence of 2.5 mM $ZnCl_2$, and this spectrum is essentially identical to that of the free cterRAP74 (B. D. Nguyen, J. G. Omichinski, and P. Legault, unpublished data). Thus, it appears that this short α-helix is either not present or unstable under our NMR conditions both in the absence and in the presence of zinc. In the crystal structure, the zinc is coordinated using three surface residues (E503, H512, and E517) located in two different protomers; thus, it was not surprising that zinc does not bind to the free cterRAP74 monomer in solution. However, the authors proposed a mechanism in which one molecule of the cterRAP74 could bind zinc, and suggested that zinc binding might be physiologically important to RAP74 function (*34*). Interestingly, preliminary in vitro binding and NMR titration experiments indicate that zinc can inhibit cterFCP binding to the cterRAP74 (B. D. Nguyen, J. G. Omichinski, and P. Legault, unpublished data). This inhibition of a FCP1−RAP74 complex by zinc may be the result of zinc binding to the cterFCP since our preliminary NMR data suggest that zinc does not bind to the cterRAP74. We are currently carrying out additional experiments to determine the molecular basis of the zinc inhibition.

*Resonance Assignments and Secondary Structure of a cterRAP74−cterFCP Complex.* It has been shown using yeast two-hybrid and in vitro binding studies that the cterRAP74 specifically interacts with the acidic carboxyl-terminal domain of FCP (cterFCP, residues 881−961). This direct interaction between RAP74 and FCP may be responsible for the RAP74 stimulation of FCP phosphatase activity.

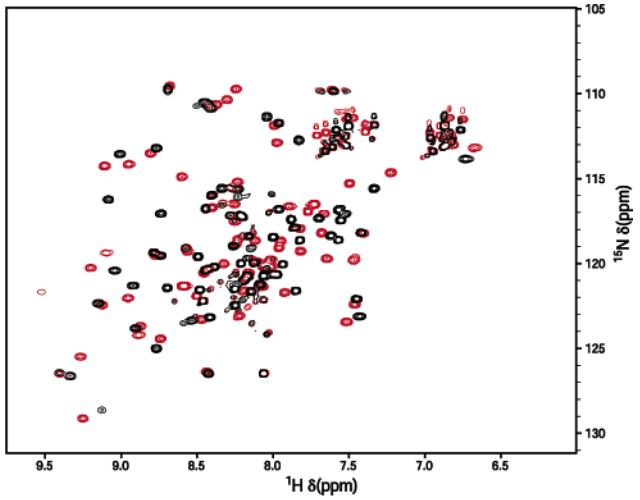


FIGURE 2: Overlay of the 2D $^{1}H-^{15}N$ HSQC spectra of the free $^{15}N$-labeled human cterRAP74 (black) and the $^{15}N$-labeled cterRAP74−unlabeled cterFCP complex (red).

To identify the binding site of the cterFCP on the cterRAP74, a RAP74−FCP complex was formed using the unlabeled cterFCP and the $^{15}N$-labeled cterRAP74. The overlay of the $^{1}H-^{15}N$ HSQC spectra of the free $^{15}N$-labeled cterRAP74 (black) and of the $^{15}N$-labeled cterRAP74−unlabeled cterFCP complex (red) is illustrated in Figure 2. Interestingly, several correlations display large changes in $^{1}H$ and $^{15}N$ chemical shifts between these two HSQC spectra. These dramatic differences in chemical shifts for several signals made it impossible to completely assign all signals in the HSQC spectrum of the cterRAP74−cterFCP complex by comparison with the signals in the HSQC spectrum of the free cterRAP74. To achieve backbone assignment of cterRAP74 in the complex, triple-resonance experiments [3D HNCO, 3D HNCACB, and 3D (HB)CBCA(CO)NNH] were carried out with a complex consisting of the $^{13}C$- and $^{15}N$-labeled cterRAP74 and the unlabeled cterFCP. A chemical shift index (CSI) analysis of $H_\alpha$, $C_\alpha$, $C_\beta$, and C′ was performed to compare the secondary structure of the free and complexed cterRAP74 (data not shown). Interestingly, this CSI analysis revealed that the secondary structure of the cterRAP74 in complex with cterFCP is virtually identical to that of the free protein (data not shown) (*35*−*37*). Given this similarity in secondary structure, it is likely that the overall three-dimensional fold of the cterRAP74 remains essentially unchanged following binding to the cterFCP. Furthermore, the majority of the amide signals in the HSQC spectra display similar $^{1}H$ and $^{15}N$ chemical shifts for the free and cterFCP-bound form of the cterRAP74 (see Figure 3 and the Discussion), providing additional evidence that the overall three-dimensional structure is not grossly altered by binding of the cterFCP.

*Mapping the Binding Site of the cterFCP on the Solution Structure of the cterRAP74.* To map the binding site for the cterFCP on the solution structure of the cterRAP74, we computed the differences in $^{1}H$, $^{13}C'$, and $^{15}N$ chemical shifts of the cterRAP74 between the free state and in complex with the cterFCP. This analysis reveals that 24 of 82 amino acid residues of the cterRAP74 displayed significant changes $\{\Delta\delta > 0.2; \Delta\delta = [(0.17\Delta N_H)^2 + (0.39\Delta C')^2 + (\Delta H_N)^2]^{1/2}\}$ upon formation of the complex (Figure 3) (*60*). These residues correspond to positions 454, 470−472, 474−482, 485, 486,

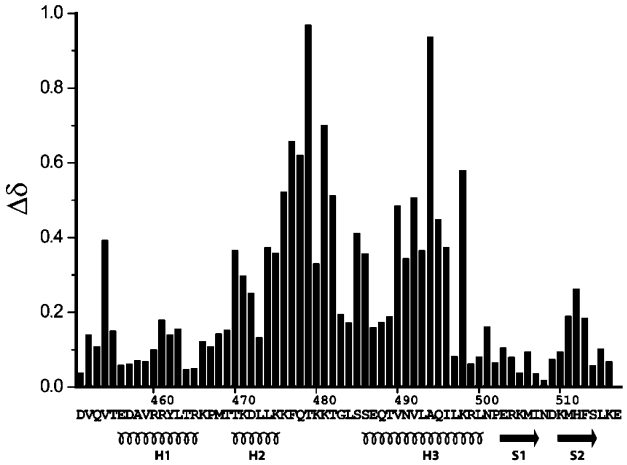


FIGURE 3: Histogram displaying the differences in chemical shifts between the cterRAP74 and the cterRAP74 in complex with the cterFCP. The differences were calculated according to the formula $\Delta\delta = [(0.17\Delta N_H)^2 + (0.39\Delta C')^2 + (\Delta H_N)^2]^{1/2}$ (*60*).

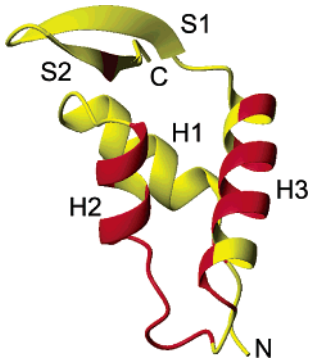


FIGURE 4: Signals which show a significant chemical shift change $\{\Delta\delta > 0.2; \Delta\delta = [(0.17\Delta N_H)^2 + (0.39\Delta C')^2 + (\Delta H_N)^2]^{1/2}\}$ (*60*) upon formation of the cterRAP74−cterFCP complex mapped onto the NMR structure of the cterRAP74 (yellow). The locations of significantly shifted signals are indicated in red.

491−496, 498, and 512 of the cterRAP74. When mapped onto the NMR solution structure of the cterRAP74, the residues displaying significant chemical shift changes are predominantly located in α-helices H2 and H3 and in the adjoining loop (L2), where these structural elements form a shallow groove on the surface of the cterRAP74 (Figure 4) (*34*). This shallow groove between H2 and H3 and adjoining loop L2 are identified here as the cterFCP-binding site of RAP74 on the basis of our NMR data. This FCP-binding site consists of several hydrophobic and basic amino acids (Figure 5) (*34*). The hydrophobic amino acids are from both the H2 and H3 α-helices, and they include L473, L474, V490, L493, and L497 (Figure 5A). The basic amino acids appear throughout the groove, and they originate from H2 (K471 and K475), H3 (K498 and R499), and loop L2 between H2 and H3 (K480 and K481) (Figure 5B). We have previously shown the importance of two basic amino acids in H2 (K471 and K475) for binding to the cterFCP by mutagenesis studies in yeast (*32*). In addition, a binding site in the hydrophobic groove between H2 and H3 had also been previously postulated by mapping these mutational data (*32*) on the crystal structure of the zinc-bound cterRAP74 (*34*). However, it is clear that residues in L2 of the cterRAP74 also play a crucial role in cterFCP binding in addition to the groove between H2 and H3.

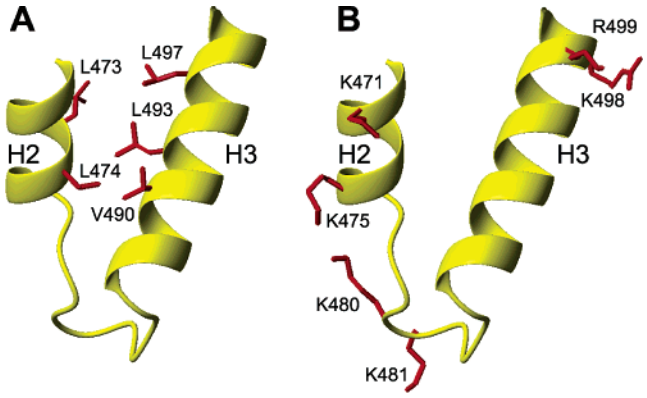


FIGURE 5: Ribbon diagram of the cterFCP-binding domain of the NMR structure of the cterRAP74. The H2 helix, the H3 helix, and the connecting loop between the H2 and H3 helices are shown in yellow. The side chains of hydrophobic (A) and basic (B) amino acids within the binding site are displayed in red.

*Mapping the Binding Site of the cterFCP on the Solution Structure of TFIIBc.* In previous studies, it has been demonstrated that the cterFCP also interacts with the first cyclin repeat of TFIIB. TFIIB has been shown to play a role in regulating the CTD phosphatase activity through its ability to block the stimulatory effect produced by the large subunit of TFIIF. We have prepared a $^{15}$N-labeled sample of a human TFIIB fragment that includes both cyclin repeats (TFIIBc, residues 102−316). The solution structure of this identical fragment of TFIIB has been previously determined by NMR spectroscopy (*33*, *61*). To identify the binding site of the cterFCP on TFIIBc, we have recorded $^1$H−$^{15}$N HSQC spectra of free and cterFCP-bound $^{15}$N-labeled TFIIBc. The $^1$H−$^{15}$N HSQC spectrum of free $^{15}$N-labeled TFIIBc is virtually identical to what has been previously reported (data not shown) (*33*, *61*). In the $^1$H−$^{15}$N HSQC spectrum of cterFCP-bound $^{15}$N-labeled TFIIBc, most of the amide signals have the same $^1$H and $^{15}$N chemical shifts as observed in the spectrum of free TFIIBc. This provides evidence that binding of the cterFCP has not significantly altered the overall fold of TFIIBc. Since complete NMR chemical shift assignments of free TFIIBc were available from the previous NMR studies (*61*), we were able to quickly identify amino acid residues of TFIIBc that are likely to form the binding site of cterFCP. From all the amide signals observed in the $^1$H−$^{15}$N HSQC spectrum of free $^{15}$N-labeled TFIIBc, there are 42 amide signals for which we detect a significant chemical shift change [$\Delta(^1H) \geq 0.03$ ppm and/or $\Delta(^{15}N) \geq 0.3$ ppm] and/or line broadening upon formation of the TFIIBc−cterFCP complex. These residue changes were mapped onto the NMR structure of TFIIBc and are located predominantly around helix D1 (residues 177−182) and helix E1 (residues 188−200) of the first cyclin repeat, where these structural elements form a shallow groove on the surface of TFIIBc (Figure 6A). This groove in the first cyclin repeat of TFIIB is identified here as the cterFCP-binding site of TFIIBc. These NMR results are consistent with in vitro binding experiments that indicated that the first cyclin repeat of yeast TFIIB is necessary and sufficient for binding to the yeast cterFCP (*32*). As was seen for the cterFCP-binding site of RAP74, the cterFCP-binding site of TFIIBc consists of several hydrophobic and several basic amino acids (Figure 7). The hydrophobic amino acids are from α-helices D1 and E1, and

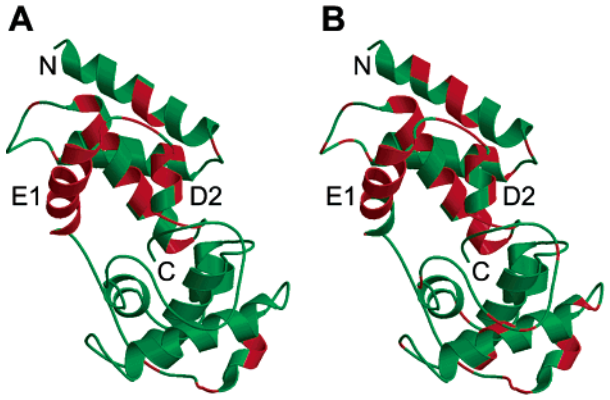


FIGURE 6: Signals which show a significant chemical shift change [$\Delta(^1H) \geq 0.03$ ppm and/or $\Delta(^{15}N) \geq 0.3$ ppm] and/or line broadening upon formation of (A) the TFIIBc−cterFCP complex and (B) the TFIIBc−VP16 activation domain complex mapped onto the NMR structure (green) of TFIIBc (*33*, *61*). The locations of shifted or broadened signals are indicated in red.

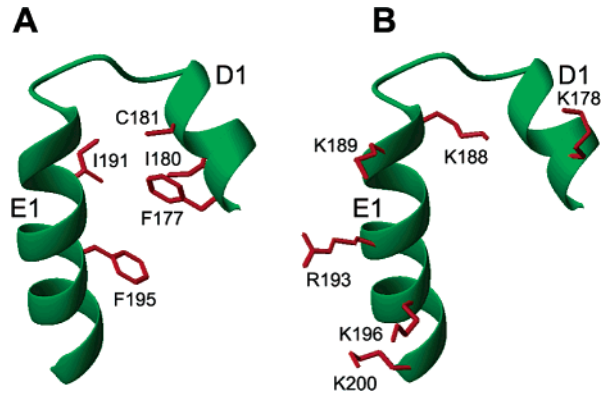


FIGURE 7: Ribbon diagram of the cterFCP-binding domain of the NMR structure of TFIIBc (*33*). The D1 and E1 helices and the connecting loop between the D1 and E1 helices are shown in green. The side chains of hydrophobic (A) and basic (B) amino acids within the binding site are displayed in red.

they include F177, I180, C181, I191, and F195 (Figure 7A). The basic amino acids appear throughout the groove, and they also originate from α-helices D1 (K178) and E1 (K188, K189, R193, K196, and K200) (Figure 7B). Again, we have previously shown the importance of two basic amino acids in helix E1 (K189 and R193) in binding to the cterFCP in mutagenesis studies in yeast (*32*).

In a previous study, the interactions of TFIIBc with the amino-terminal zinc ribbon of TFIIB and the activation domain of the herpes simplex virus protein VP16 have been mapped. This study used the same TFIIBc protein that was used in our study and NMR methods similar to those described here (Figure 6B) (*61*). To identify residues forming the cterFCP-binding site of TFIIBc, we applied the same criteria of chemical shift changes and line broadening that had been previously used to identify residues forming the amino-terminal TFIIB-binding site and the VP16-binding site of TFIIBc (*61*). Interestingly, we find a remarkable similarity in the residues of the first cyclin repeat of TFIIBc that display chemical shift changes following addition of the cterFCP, the amino-terminal zinc ribbon of TFIIB and the activation domain of VP16 (Figure 6B) (*61*).

*Comparison of cterFCP-Binding Sites for RAP74 and TFIIBc.* From in vitro binding experiments, it has been shown that the cterFCP directly interacts with RAP74 and TFIIB,

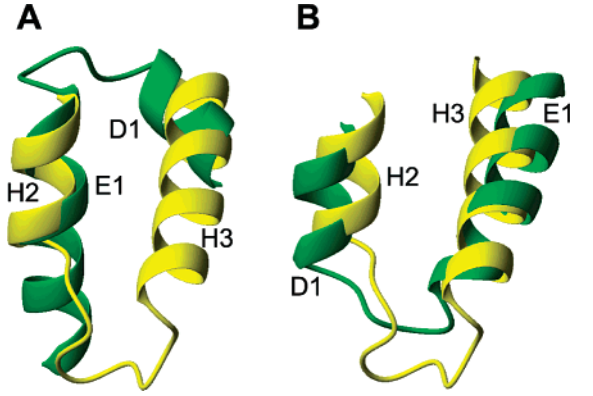


FIGURE 8: Ribbon diagram of two backbone superpositions of the cterFCP-binding domains of the cterRAP74 (yellow) and TFIIBc (green) (see the text for details).

and that this binding involves a homologous stretch of eight amino acids found in the cterRAP74 (residues 468−475, helix H2) and in the first cyclin repeat of TFIIBc (residues 186−193, helix E1) (*32*). The NMR studies presented here support these in vitro binding results, and they help us to more thoroughly map the cterFCP-binding site of both RAP74 and hTFIIB. In both RAP74 and TFIIB, the cterFCP-binding site consists of two α-helices and a connecting loop (Figures 4 and 6), which form a shallow groove that contains several hydrophobic and several basic amino acids (Figures 5 and 7). The binding groove in RAP74 is between α-helices H2 and H3, while the binding groove in TFIIB is between α-helices D1 and E1. The homologous stretch of eight amino acids identified from the in vitro binding experiments is located in helix H2 in RAP74 and in helix E1 in TFIIB. To compare these two cterFCP-binding sites, the FCP-binding site of cterRAP74 was superimposed onto the FCP-binding site of TFIIBc. Two backbone superpositions were performed to find an optimal alignment of these two cterFCP-binding sites (Figure 8). In the first superposition (Figure 8A), backbone residues of helices H2 (residues 470−476) and H3 (residues 495−500) of cterRAP74 were superimposed onto backbone residues of helices E1 (residues 188−194) and D1 (residues 177−182) of TFIIBc, respectively, and an rmsd of 2.8 Å was obtained for these atoms. In this superposition, the homologous stretches of amino acids found in helix H2 of the cterRAP74 (T470, K471, D472, L473, L474, K475, and F476) and helix E1 of human TFIIBc (K188, K189, E190, I191, G192, R193, and C194) are aligned residue by residue in three dimensions according to their sequence alignment (*32*). This superposition places the connecting loops between the two helices at opposite ends of the grooves. In the second superposition (Figure 8B), backbone residues of helices H2 (residues 470−476) and H3 (residues 486−498) of the cterRAP74 were superimposed onto helices D1 (residues 176−182) and E1 (residues 188−200) of TFIIBc, respectively, and an rmsd of 2.4 Å was obtained for these atoms. In this superposition, the two longer helices (H3 and E1) are aligned, the two shorter helices are aligned (H2 and D1), and the connecting loops between the two helices are found at the same end of the grooves. However, in this superposition, the homologous stretches of eight amino acids for the cterRAP74 and TFIIBc are found on opposite faces of the grooves. Interestingly, it is clear that the secondary structure elements are better aligned in

the second superposition, although their homologous stretches of residues do not overlap in three dimensions.

## DISCUSSION

The phosphorylation state of the CTD of RNAP II plays a crucial role in regulating transcriptional elongation. Following the cloning of the CTD-specific phosphatase FCP, in vitro binding studies have demonstrated that FCP directly interacts with distinct domains within the general transcription factors TFIIF and TFIIB. Specifically, it was shown that residues 436−517 of RAP74 (cterRAP74) are sufficient and necessary for interaction with residues 881−961 of human FCP (cterFCP) (*24*, *26*). In addition, it has been shown that the first cyclin repeat of TFIIB interacts with two regions of yeast FCP, and one of these two regions corresponds to the human cterFCP (*32*). With this paper, we start to structurally characterize the interaction of FCP with these two general transcription factors.

In the initial phase of the project, we determined the solution structure of the cterRAP74 in the free state using homonuclear and heteronuclear NMR spectroscopy (Figure 1A). The solution structure of the free cterRAP74 is very similar to the X-ray crystal structure of a zinc-bound cterRAP74 fragment (*34*), and a backbone (N, $C_\alpha$, and C′) superposition of the two structures yielded an atomic rmsd of 0.8 Å (Figure 1B). In the X-ray studies with the zinc-bound cterRAP74, the overall fold of the cterRAP74 was found to be very similar to those of several eukaryotic winged-helix domains except for the presence of a short four-residue α-helix found within the loop region (L2) connecting α-helices H2 and H3 (*34*). There is no evidence for this short α-helix in the NMR structure of the free cterRAP74, and from our analysis of chemical shifts and NOE data of free cterRAP74. It is possible that formation of this short α-helix in the crystal structure results from zinc binding. Initially, this would appear rather unlikely since the zinc binding in the crystal structure bridges two adjacent molecules and involves amino acids remote from the short α-helix. To test the role of zinc binding, we compared the $^1H-^{15}N$ HSQC spectrum of the free cterRAP74 with that of the cterRAP74 in the presence of 2.5 mM $ZnCl_2$. If formation of the small α-helix in L2 were zinc-dependent or if zinc bound specifically to the cterRAP74, we would observe chemical shift differences for at least a small number of signals between these two HSQC spectra. However, these two spectra are essentially identical (B. D. Nguyen, J. G. Omichinski, and P. Legault, unpublished data), which indicates that, under these conditions, zinc does not induce a stable structural change in the cterRAP74. Interestingly, preliminary in vitro binding and NMR studies indicate that zinc ions can actually inhibit binding of the cterFCP to the cterRAP74 (B. D. Nguyen, J. G. Omichinski, and P. Legault, unpublished data). Determination of the exact physiological importance of this observation will require further testing.

An alternative explanation for the formation of the short α-helix (residues 479−482) is crystal packing. Close examination of the crystal structure reveals two hydrogen bonds that likely stabilize this short α-helix. One hydrogen bond is between the guanidinium group of R499 on one protomer and the backbone carbonyl oxygen of T482 on a second protomer. The second hydrogen bond is between the ammonium group of K481 and the backbone carbonyl oxygen of G448 on the same protomer. This later interaction involves a residue (G448) that is not present in the wild-type RAP74 sequence. Four amino acids (G445, P446, L447, and G448) originate form the pGEX vector used for protein expression in the crystallography study, and the corresponding residues in our NMR study are wild-type residues (T445, T446, P447, and N448). It is therefore not entirely clear if the short α-helix is caused by zinc binding, crystal packing, non-native amino acid residues in the protein, or a combination of the above. However, it appears that this α-helix is unique to the crystal structure.

Following the completion of the cterRAP74 structure, NMR studies were conducted with a cterRAP74−cterFCP complex to determine the cterFCP-binding site on the cterRAP74. On the basis of changes in chemical shift observed between the $^1H-^{15}N$ HSQC NMR spectra of the cterRAP74 and the cterRAP74−cterFCP complex (Figures 2 and 3), we concluded that the FCP-binding site is located in the shallow groove formed between α-helices H2 and H3 in the cterRAP74 (Figure 4). Interestingly, the cterFCP-binding site of the cterRAP74 contains a rather large number of hydrophobic and basic amino acids (Figure 5). The hydrophobic amino acids are L473 and L474 of α-helix H2 and V490, L493, and L497 of α-helix H3 (Figure 5A). The basic amino acids in the cterFCP-binding site originate from H2 (K471 and K475), H3 (K498 and R499), and loop L2 between H2 and H3 (K480 and K481) of the cterRAP74 (Figure 5B). Given the large number of acidic and hydrophobic residues in the cterFCP, the presence of a large number of basic and hydrophobic amino acids in the cterFCP-binding site of the cterRAP74 is not surprising. In fact, this binding site had also been previously postulated on the basis of mutational data (*32*) mapped on the crystal structure of the zinc-bound cterRAP74 (*34*). The precise interactions of the cterRAP74 with the cterFCP most likely involve important ionic and hydrophobic interactions.

To identify the cterFCP-binding site of TFIIBc, NMR studies were conducted with a TFIIBc−cterFCP complex. On the basis of changes in chemical shift observed between the $^1H-^{15}N$ HSQC NMR spectra of the free TFIIBc and the TFIIBc−cterFCP complex, we concluded that the FCP-binding site is located in the shallow groove formed between α-helices D1 and E1 of TFIIBc (Figure 6A). Consistent with the in vitro binding data, we were not able to observe any significant chemical shift changes in amino acids located in the second cyclin repeat in the TFIIBc−cterFCP complex (*32*). Interestingly, the carboxyl-terminal activation domain of the herpes simplex virus VP16 and the amino-terminal zinc ribbon of TFIIB also bind to the same region (*61*) of the first cyclin repeat of TFIIB (Figure 6). Like the cterFCP, the activation domain of VP16 (residues 412−490) also contains a large number of acidic and hydrophobic residues that are important for binding to several general transcription factors (*62*). Thus, the binding of the VP16 activation domain and the cterFCP to the first cyclin repeat of TFIIB appears to be highly analogous.

To compare the cterFCP-binding sites of the cterRAP74 and TFIIBc, we superimposed their backbones in two different orientations (Figure 8). In the first superposition (Figure 8A), the H2 and H3 helices of the cterRAP74 were superimposed onto the E1 and D1 helices of TFIIBc,

respectively. Although the homologous stretch of eight amino acids (helix H2 of the cterRAP74 and helix E1 of TFIIBc) are perfectly aligned, the helices and connecting loops are not. A superior alignment of the secondary structure elements is achieved in the second superposition (Figure 8B), where backbone residues of helices H2 and H3 of the cterRAP74 were superimposed with helices D1 and E1 of TFIIBc, respectively. However, in this second superposition, the homologous stretches of eight amino acids are found on opposite faces of the shallow groove. On the basis of rmsd values and a superior secondary structural alignment, it appears that the cterFCP-binding sites of the cterRAP74 and TFIIBc are aligned optimally in this second superposition. However, if we consider that the acidic-rich cterFCP must make critical contacts with basic amino acids of the cterRAP74 and TFIIBc, the first superposition represents the appropriate superposition of the binding sites. Interestingly, this superposition gives a similar three-dimensional spatial distribution of the basic amino acids present in the two binding sites, and the conserved basic residues (K471 and K475 in RAP74 and K189 and R193 in TFIIB) of the homologous stretches of eight amino acids are aligned perfectly in this superposition (Figure 8).

## CONCLUSIONS

In this paper, we have determined the NMR solution structure of the free human cterRAP74, and we have used NMR methods to map the cterFCP-binding sites for both the cterRAP74 and human TFIIB. The cterFCP-binding sites of both RAP74 and TFIIBc form shallow grooves on the protein surface, and they are rich in hydrophobic and positively charged amino acid side chains. The cterFCP-binding site of TFIIB is very similar to the previously characterized TFIIB-binding site for the acidic transcriptional activator protein VP16 (*61*). Like the VP16 activation domain, the cterFCP is highly acidic and has been shown to be a transcriptional activator when artificially recruited to a promoter (*31*). Preliminary results suggest that the free cterFCP is mostly disordered based on the chemical shift dispersion in a 1D $^1$H NMR experiment (B. D. Nguyen, J. G. Omichinski, and P. Legault, data not shown). The dramatic changes in chemical shifts observed for the cterRAP74 following the interaction with the cterFCP (Figure 3) indicate that cterFCP may adopt a more defined structure upon binding RAP74. NMR studies have shown that the acidic activation domain of VP16 (residues 412−490) is also disordered in the free form (*63*), but a 17-amino acid fragment of this domain (residues 469−485) adopts a more ordered α-helical conformation following binding with the TBP-associated factor hTAF$_{II}$31 (*64*). Like VP16 and other transactivation domains (*64*−*67*), the cterFCP may also adopt an α-helical conformation upon binding to its target cterRAP74. We are currently carrying out additional NMR experiments to obtain a three-dimensional structure of the cterRAP74−cterFCP complex.

## ACKNOWLEDGMENT

We acknowledge Brian Bissell for assistance in preparation of protein samples. We thank Dr. Danny Reinberg for the clone expressing TFIIBc, Dr. Jacques Archambault for RAP74 and FCP clones, Dr. Frank Delaglio and Dr. Daniel Garrett for supplying NMR data processing and analysis programs, Dr. Stephan Grzesiek for help in some of the preliminary experiments, and Dr. Lewis Kay and Dr. Wing-Yiu Choy for supplying NMR pulse sequences and a modified version of CNS.

## SUPPORTING INFORMATION AVAILABLE

$^{15}$N−$^1$H heteronuclear NOE data (*67*) recorded on the $^{15}$N-labeled cterRAP74 at pH 6.5 and 27 °C. This material is available free of charge via the Internet at http://pubs.acs.org.

## REFERENCES

1. Archambault, J., and Friesen, J. D. (1993) *Microbiol. Mol. Biol. Rev. 57*, 703−724.
2. Hampsey, M. (1998) *Microbiol. Mol. Biol. Rev. 62*, 465−503.
3. Zawel, L., and Reinberg, D. (1993) *Prog. Nucleic Acid Res. Mol. Biol. 44*, 67−108.
4. Ranish, J. A., Yudkovsky, N., and Hahn, S. (1999) *Genes Dev. 13*, 49−63.
5. Greenleaf, A. L. (1993) *Trends Biochem. Sci. 18*, 117−119.
6. Buratowski, S., Hahn, S., Guarente, L., and Sharp, P. A. (1989) *Cell 56*, 549−561.
7. Allison, L. A., Moyle, M., Shales, M., and Ingles, C. J. (1985) *Cell 42*, 599−610.
8. Corden, J. L., Cadena, D. L., Ahearn, J. M., and Dahmus, M. E. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 7934−7938.
9. McCracken, S., Fong, N., Yankulov, K., Bothers, G., Siderovski, D., Hessel, A., Forster, S., Program, A. E., Shuman, S., and Bentley, D. L. (1997) *Genes Dev. 11*, 3306−3318.
10. McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G. H., Greenblatt, J., Patterson, S. D., Wickens, M., and Bentley, D. L. (1997) *Nature 385*, 357−361.
11. Cadena, D. L., and Dahmus, M. E. (1987) *J. Biol. Chem. 262*, 12468−12474.
12. Lu, H., Flores, O., Weinmann, R., and Reinberg, D. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 10004−10008.
13. O'Brien, T. S., Hardin, S., Greenleaf, A., and Lis, J. T. (1994) *Nature 370*, 75−77.
14. Lu, H., Zawel, L., Fisher, L., Egly, J. M., and Reinberg, D. (1992) *Nature 358*, 641−645.
15. Feaver, W. J., Gileadi, O., Li, Y., and Kornberg, R. D. (1991) *Cell 67*, 1223−1230.
16. Serizawa, H. R., Conaway, R. C., and Conaway, J. W. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 7476−7480.
17. Marshall, N. F., Peng, J., Xie, Z., and Price, D. H. (1996) *J. Biol. Chem. 271*, 27176−27183.
18. Cujec, T. P., Okamoto, H., Fujinaga, K., Meyer, J., Chamberlin, H., Morgan, D. O., and Peterlin, B. M. (1997) *Genes Dev. 11*, 2645−2657.
19. Young, R. A. (1991) *Annu. Rev. Biochem. 60*, 689−715.
20. Dahmus, M. E. (1996) *J. Biol. Chem. 271*, 19009−19012.
21. Chambers, R. S., and Dahmus, M. E. (1994) *J. Biol. Chem. 269*, 26243−26248.
22. Chambers, R. S., and Dahmus, M. E. (1996) *J. Biol. Chem. 271*, 24498−24504.
23. Chambers, R. S., Wang, B. Q., Burton, Z. F., and Dahmus, M. E. (1995) *J. Biol. Chem. 270*, 14962−14969.
24. Archambault, J., Pan, G., Dahmus, G. K., Cartier, M., Marshall, N., Zhang, S., Dahmus, M. E., and Greenblatt, J. (1998) *J. Biol. Chem. 273*, 27593−27601.
25. Cho, H., Kim, T.-K., Mancebo, H., Lane, W. S., Flores, O., and Reinberg, D. (1999) *Genes Dev. 13*, 1540−1552.
26. Archambault, J., Chambers, R. S., Kobor, M. S., Ho, Y., Cartier, M., Bolotin, D., Andrews, B., Kane, C. M., and Greenblatt, J. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 14300−14305.
27. Collet, J. F., Stroobant, V., Pirard, M., Delpierre, G., and Schaftingen, E. V. (1998) *J. Biol. Chem. 273*, 14107−14112.
28. Thaller, M. C., Schippa, S., and Rossoline, G. M. (1998) *Protein Sci. 7*, 1647−1652.
29. Kobor, M. S., Archambault, J., Lester, W., Holstege, F. C. P., Gileadi, O., Jansman, D. B., Jennings, E. G., Kouyoumdjian, F., Davidson, A. R., Young, R. A., and Greenblatt, J. (1999) *Mol. Cell 4*, 55−62.
30. Bork, P., Hofmann, K., Bucher, P., Neuwald, A. F., Altschul, S., and Koonin, E. V. (1997) *FASEB J. 11*, 68−76.

31. Licciardo, P., Ruggiero, L., Lania, L., and Majello, B. (2001) *Nucleic Acids Res. 29*, 3539−3545.
32. Kobor, M. S., Simon, L. S., Omichinski, J. G., Zhong, G., Archambault, J., and Greenblatt, J. (2000) *Mol. Cell. Biol. 20*, 7438−7449.
33. Bagby, S., Kim, S., Maldonado, E., Tong, K. I., Reinberg, D., and Ikura, M. (1995) *Cell 82*, 857−867.
34. Kamada, K., Angelis, J. D., Roeder, R. G., and Burley, S. K. (2001) *Proc. Natl. Acad. Sci. U.S.A. 98*, 3115−3120.
35. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1992) *Biochemistry 31*, 1647−1651.
36. Wishart, D. S., and Sykes, B. D. (1994) *J. Biomol. NMR 4*, 171−180.
37. Wishart, D. S., and Nip, A. M. (1998) *Biochem. Cell. Biol. 76*, 153−163.
38. Ha, I., Roberts, S., Maldonado, E., Sun, X., Kim, L.-U., Green, M., and Reinberg, D. (1993) *Genes Dev. 7*, 1021−1032.
39. Kay, L. E., Keifer, P., and Saarinen, T. (1992) *J. Am. Chem. Soc. 114*, 10663−10665.
40. Marion, D., Kay, L. E., Sparks, S. W., Torchia, D. A., and Bax, A. (1989) *J. Am. Chem. Soc. 111*, 1515−1517.
41. Wittekind, M., and Mueller, L. (1993) *J. Magn. Reson., Ser. B 101*, 201−205.
42. Grzesiek, S., Doebeli, H., Gentz, R., Garotta, G., Labhardt, A. M., and Bax, A. (1992) *Biochemistry 31*, 8180−8190.
43. Muhandiram, D. R., and Kay, L. E. (1994) *J. Magn. Reson., Ser. B 103*, 203−216.
44. Grzesiek, S., Ikura, M., Clore, G. M., Gronenborn, A. M., and Bax, A. (1992) *J. Magn. Reson. 96*, 215−221.
45. Grzesiek, S., Anglister, J., and Bax, A. (1993) *J. Magn. Reson. 101*, 114−121.
46. Zuiderweg, E. R. P., McIntosh, L. P., Dahlquist, F. W., and Fesik, S. W. (1990) *J. Magn. Reson. 86*, 210−216.
47. Bax, A., Sparks, S. W., and Torchia, D. A. (1989) *Methods Enzymol. 176*, 134−150.
48. Pascal, S. M., Muhandiram, D. R., Yamazaki, T., Forman-Kay, J. D., and Kay, L. E. (1994) *J. Magn. Reson. 103*, 197−201.
49. Wishart, D. S., Bigam, C. G., Yao, J., Dyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1995) *J. Biomol. NMR 6*, 135−140.
50. Markley, J. L., Bax, A., Arata, Y., Hilbers, C. W., and Wuthrich, K. (1998) *Pure Appl. Chem. 70*, 117−142.
51. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR 6*, 277−293.
52. Garrett, D. S., Powers, R., Gronenborn, A. M., and Clore, G. M. (1991) *J. Magn. Reson. 95*, 214−220.
53. Cornilescu, G., Delaglio, F., and Bax, A. (1999) *J. Biomol. NMR 13*, 289−302.
54. Brunger, A. T., Adams, P. D., Clore, G. M., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D54*, 905−921.
55. Kuszewski, J., Gronenborn, A. M., and Clore, G. M. (1996) *Protein Sci. 5*, 1067−1080.
56. Choy, W.-Y., Tollinger, M., Mueller, G. A., and Kay, L. E. (2001) *J. Biomol. NMR 21*, 31−40.
57. Laskowski, R. A., Antoon, J., Rullmann, C., Macarthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR 8*, 477−486.
58. Koradi, R., Billeter, M., and Wüthrich, K. (1996) *J. Mol. Graphics 14*, 51−55.
59. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
60. Farmer, B. T., Constantine, K. L., Goldfarb, V., Friedrichs, M. S., Wittekind, M., Yanchunas, J., Robertson, J. G., and Mueller, L. (1996) *Nat. Struct. Biol. 3*, 995−997.
61. Hayashi, F., Ishima, R., Liu, D., Tong, K. I., Kim, S., Reinberg, D., Bagby, S., and Ikura, M. (1998) *Biochemistry 37*, 7941−7951.
62. Xiao, H., Pearson, A., Coulombe, B., Truant, R., Zhang, S., Regier, J. L., Trienzenberg, S. J., Reinberg, D., Flores, O., Ingles, C. J., and Greenblatt, J. (1994) *Mol. Cell. Biol. 14*, 7013−7024.
63. O'Hare, P., and Williams, G. (1992) *Biochemistry 31*, 4150−4156.
64. Uesugi, M., Nyanguile, O., Lu, H., Levine, A. J., and Verdine, G. L. (1997) *Science 277*, 1310−1313.
65. Kussie, P. H., Gorina, S., Marechal, V., Elenbaas, B., Moreau, J., Levine, A. J., and Pavletich, N. P. (1996) *Science 274*, 948−953.
66. Dames, S. A., Martinez-Yamount, M., Guzman, R. N. D., Dyson, H. J., and Wright, P. E. (2002) *Proc. Natl. Acad. Sci. U.S.A. 99*, 5271−5276.
67. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) *Biochemistry 33*, 5984−6003.

BI0265473